WHAT IS CLAIMED IS:

1. A computer-implemented method for supervised artificial neural network machine learning, comprising:

reducing the dimensionality of the received data to enhance machine learning performance based on the dimensionality;

specifying the supervised neural network architecture;

initializing weights to establish connection strengths between the received data and predicted values;

performing supervised machine learning using the specified architecture, initialized weights, and the received data including the reduced dimensionality to predict values; and

revising the initialized weights of the network based on a normalized system error threshold value to generate a learnt neural network having a reduced error in weight space.

2. The method of claim 1, wherein the data comprises:

data selected from the group consisting of static data and real-time data.

3. The method of claim 2, wherein reducing the dimensionality of the received data to enhance machine learning performance based on the dimensionality, further comprises:

receiving data;

checking dimensionality of the received data; and

reducing the dimensionality of the received data to enhance machine learning performance based on the outcome of the checking.

4. The method of claim 3, wherein checking the dimensionality of the received data to improve performance during the machine learning further comprises:

comparing the dimensionality of the received data to a threshold value; and

reducing the dimensionality of the received data to increase machine learning performance based on the outcome of the comparison.

5. The method of claim 4, wherein reducing the dimensionality of the received data comprises;

reducing the number of attributes in the received data using Principal Component Analysis.

6. The method of claim 4, wherein comparing the dimensionality of the received data to the threshold value comprises:

comparing the number of attributes in the received data to the threshold value.

7. The method of claim 6, wherein the threshold value comprises:

greater than or equal to 25 attributes.

8. The method of claim 2, further comprising:

if the received data is static data, then reading a sample of the received static data using a predetermined window length; and

if the received data is real-time data, then reading a sample of the received real-time data using a dynamically varying window of predetermined window length.

9. The method of claim 8, further comprising:

if the received data is real-time data, then repeating the receiving of the real-time data using a dynamically varying window of predetermined window length.

10. The method of claim 2, wherein specifying the supervised neural network architecture comprises:

specifying the learning parameters for the neural network selected from the group consisting of number of input nodes, number of hidden layers, number of nodes in each of the layers, number of nodes at the output layer, and learning rate; and

allocating dynamic storage for updating the initialized weights and to store the trend between input and output nodes during each iteration using the specified neural network architecture.

11. The method of claim 2, wherein initializing the weights further comprises: initializing the weights using random weights.

12. The method of claim 2, wherein revising the initialized weights of the network based on a normalized system error threshold value, further comprises:

comparing the predicted values to a normalized system error threshold value; and

revising the initialized weights of the neural network based on the outcome of the comparison to generate a learnt neural network having a reduced error in weight space.

13. The method of claim 12, wherein comparing the predicted values to the normalized system error threshold value further comprises:

computing the normalized system error by using desired values and the predicted values to reduce error in the weight space using a gradient descent technique; and

comparing the computed normalized system error with the normalized system error threshold value.

14. The method of claim 13, further comprising:

repeating the performing and comparing steps until the computed normalized difference error is less than or equal to the normalized system error threshold value.

15. The method of claim 1, further comprising:

using a Hessian matrix to enhance the learning rate of the neural network; and

using a function approximation neighborhood technique to perturb the learning parameters of the neural network to further enhance the learning rate of the neural network.

16. The method of claim 1, further comprising:

validating the learnt neural network to verify the reliability of the learnt neural network.

17. The method of claim 16, wherein validating the neural network further comprises:

performing supervised learning using the learnt neural network to predict the values;

computing accuracy of the predicted values by comparing the predicted values with the known values;

comparing the computed accuracy with an accuracy value; and

repeating the varying learning rate, performing, comparing, and validating steps based on the outcome of the comparison to further enhance the reliability of the learnt neural network.

18. The method of claim 16, further comprising:

inputting unknown values into the validated neural network; and

predicting the values by performing supervised learning on the validated neural network.

19. A computer readable medium having computer-executable instructions for supervised artificial neural network learning, comprising:

receiving data;

checking dimensionality of the received data;

reducing the dimensionality of the received data to enhance machine learning performance based on the outcome of the checking;

specifying the supervised neural network architecture;

initializing weights to establish connection strengths between the received data and predicted values;

performing supervised machine learning using the specified architecture, initialized weights, and the received data including the reduced dimensionality to predict values;

comparing the predicted values to a normalized system error threshold value; and

revising the initialized weights of the neural network based on the outcome of the comparison to generate a learnt neural network having reduced error in weight space.

20. The computer readable medium of claim 19, wherein the data comprises:

data selected from the group consisting of static data and real-time data.

21. The computer readable medium of claim 20, wherein checking the dimensionality of the received data to improve performance during the machine learning further comprises:

comparing the dimensionality of the received data to a threshold value; and

reducing the dimensionality of the received data to increase machine learning performance based on the outcome of the comparison.

22. The computer readable medium of claim 21, wherein reducing the dimensionality of the received data comprises;

reducing the number of attributes in the received data using Principal Component Analysis.

23. The computer readable medium of claim 21, wherein comparing the dimensionality of the received data to the threshold value comprises:

comparing the number of attributes in the received data to the threshold value.

24. The computer readable medium of claim 23, wherein the threshold value comprises:

greater than or equal to 25 attributes.

25. The computer readable medium of claim 20, further comprising:

if the received data is static data, then reading a sample of the received static data using a predetermined window length; and

if the received data is real-time data, then reading a sample of the received real-time data using a dynamically varying window of predetermined window length.

26. The computer readable medium of claim 25, further comprising:

if the received data is real-time data, then repeating the receiving of the real-time data using a dynamically varying window of predetermined window length.

27. The computer readable medium of claim 20, wherein specifying the supervised neural network architecture comprises:

specifying the learning parameters for the neural network selected from the group consisting of number of input nodes, number of hidden layers, number of nodes in each of the layers, number of nodes at the output layer, and learning rate; and

allocating dynamic storage for updating the initialized weights and to store the trend between input and output nodes during each iteration using the specified neural network architecture.

28. The computer readable medium of claim 20, wherein initializing the weights further comprises:

initializing the weights using random weights.

29. The computer readable medium of claim 20, wherein comparing the predicted values to the normalized system error threshold value further comprises:

computing the normalized system error by using desired values and the predicted values to reduce error in the weight space using gradient descent technique; and

comparing the computed normalized system error with the normalized system error threshold value.

30. The computer readable medium of claim 29, further comprising:

repeating the performing and comparing steps until the computed normalized difference error is less than or equal to the normalized system error threshold value.

31. The computer readable medium of claim 19, further comprising:

using a Hessian matrix to enhance learning rate of the neural network; and

using a function approximation neighborhood technique to perturb the learning parameters of the neural network to further enhance the learning rate of the neural network.

32. The computer readable medium of claim 19, further comprising:

validating the learnt neural network to verify the reliability of the learnt neural network.

33. The computer readable medium of claim 32, wherein validating the neural network further comprises:

performing supervised learning using the learnt neural network to predict the values;

computing accuracy of the predicted values by comparing the predicted values with the known values;

comparing the computed accuracy with an accuracy value; and

repeating the varying learning rate, performing, and comparing steps based on the outcome of the comparison.

34. The computer readable medium of claim 32, further comprising:

using the validated neural network for predicting the values.

35. A computer system for machine learning in a sparse data environment, comprising:

a storage device;

an output device; and

a processor programmed to repeatedly perform a method, comprising:

receiving the data;

checking dimensionality of the received data;

reducing the dimensionality of the received data to enhance machine learning performance based on the outcome of the checking;

specifying the supervised neural network architecture;

initializing weights to establish connection strengths between the received data and predicted values;

performing supervised machine learning using the specified architecture, initialized weights, and the received data including the reduced dimensionality to predict the values;

comparing the predicted values to a normalized system error threshold value; and

revising the initialized weights of the neural network based on the outcome of the comparison to generate a learnt neural network having reduced error in weight space.

36. The system of claim 35, wherein the data comprises:

data selected from the group consisting of static data and real-time data.

37. The system of claim 36, wherein checking the dimensionality of the received data to improve performance during the machine learning further comprises:

comparing the dimensionality of the received data to a threshold value; and

reducing the dimensionality of the received data to increase machine learning performance based on the outcome of the comparison.

38. The system of claim 37, wherein reducing the dimensionality of the received data comprises:

reducing number of attributes in the received data using Principal Component Analysis.

39. The system of claim 37, wherein comparing the dimensionality of the received data to the threshold value comprises:

comparing the number of attributes in the received data to the threshold value.

40. The system of claim 39, wherein the threshold value comprises:

greater than or equal to 25 attributes.

41. The system of claim 36, further comprising:

if the received data is static data, then reading a sample of the received static data using a predetermined window length; and

if the received data is real-time data, then reading a sample of the received real-time data using a dynamically varying window of predetermined window length.

42. The system of claim 41, further comprising:

if the received data is real-time data, then repeating the receiving of the real-time data using a dynamically varying window of predetermined window length.

43. The system of claim 36, wherein specifying the supervised neural network architecture comprises:

specifying the learning parameters for the neural network selected from the group consisting of number of input nodes, number of hidden layers, number of nodes in each of the layers, number of nodes at the output layer, and learning rate; and

allocating dynamic storage for updating the initialized weights and to store the trend between input and output nodes during each iteration using the specified neural network architecture.

44. The system of claim 36, wherein initializing the weights further comprises: initializing the weights using random weights.

45. The system of claim 36, wherein comparing the predicted values to the normalized system error threshold value further comprises:

computing the normalized system error by using desired values and the predicted values to reduce error in the weight space using gradient descent technique; and

comparing the computed normalized system error with the normalized system error threshold value.

46. The system of claim 45, further comprising:

repeating the performing and comparing steps until the computed normalized difference error is less than or equal to the normalized system error threshold value.

47. The system of claim 35, further comprising:

using a Hessian matrix to enhance learning rate of the neural network; and

using a function approximation neighborhood technique to perturb the learning parameters of the neural network to further enhance the learning rate of the neural network.

48. The system of claim 35, further comprising:

validating the learnt neural network to verify the reliability of the learnt neural network.

49. The system of claim 48, wherein validating the neural network further comprises:

performing supervised learning using the learnt neural network to predict values;

computing accuracy of the predicted values by comparing the predicted values with the known values;

comparing the computed accuracy with an accuracy value; and

repeating the performing and comparing steps based on the outcome of the comparison.

50. The system of claim 48, further comprising:

using the validated neural network for predicting the values.

51. A computer-implemented system for supervised artificial neural network learning, comprising:

a receive module to receive data;

a reading module to read the received data;

an analyzer to check dimensionality of the read data and reduce the dimensionality of the read data to enhance machine learning performance based on the outcome of the checking;

wherein the analyzer specifies neural network architecture and initializes weights to establish connection strengths between the read data and predicted values obtained using the neural network, wherein the analyzer performs supervised learning using the specified architecture, initialized weights, and the read data including the reduced dimensionality to predict the values; and

a comparator coupled to the analyzer, compares the predicted values to a normalized system error threshold value, wherein the analyzer revises the initialized weights of the neural network based on the outcome of the comparison to generate a learnt neural network having a reduced error in weight space.

52. The system of claim 51, further comprising:

a database coupled to the receive module to receive and store data.

53. The system of claim 52, wherein the data comprises:

data selected from the group consisting of static data and real-time data.

54. The system of claim 53, wherein the comparator compares the dimensionality of the read data to a threshold value, and the analyzer reduces the dimensionality of the read data to increase neural network learning performance based on the outcome of the comparison by the comparator.

55. The system of claim 54, wherein the threshold value is greater than or equal to 25.

56. The system of claim 55, wherein the comparator compares number of attributes in the read data to the threshold value.

57. The system of claim 54, wherein the analyzer reduces the dimensionality of the read data by reducing the number of attributes in the read data using Principal Component Analysis.

58. The system of claim 57, wherein the analyzer reduces the dimensionality of the read data by forming a Covariance Matrix using the equation:

$$C_{n \times n} = X^t * X$$

wherein the received data is inputted in a matrix form (say $X_{m \times n}$);

and eigen values and eigen vectors are computed from the formed Covariance Matrix using the equation:

$$(C - \lambda I)U_i = 0 \rightarrow \quad (1)$$

wherein $\lambda = (\lambda_1, \lambda_2, \ldots.., \lambda_n)$ are the roots of the equation, solving the equation (1) gives eigen values, and $U_i = (u_{i1}, u_{i2}, \ldots., u_{in})$ gives the corresponding eigen vectors;

and principal components are selected using the equation:

$$\left\{1,2,..k \,\middle|\, \sum_{i=1}^{k}\lambda_i \Big/ \sum_{i=1}^{n}\lambda_i \geq \beta\right\}$$

where β is the cutoff percentage (~85%);

and features are further selected using the equation:

$$\left\{ j \,\middle|\, U_{ij} \geq \max_{j}\left\{U_{ij}\right\} \Big/ 2 \,,\, i \in \text{cuttoff } (k) \right\}$$

to reduce the dimensionality of the received data.

59. The system of claim 54, wherein the reading module reads a sample of the received data using a predetermined window length.

60. The system of claim 59, wherein the reading module reads the sample of the received data using a predetermined window length when the read data is static data, and the reading module reads a sample of the received data using a dynamically varying window of predetermined length when the read data is real-time data.

61. The system of claim 60, wherein the reading module repeats the reading of the sample of the received data using a dynamically varying window of predetermined window length when the received data is real-time data.

62. The system of claim 51, wherein the analyzer specifies the learning parameters for the neural network using the learning parameters selected from the group consisting of number of input nodes, number of hidden layers, number of nodes in each of the layers, number of nodes at the output layer, learning rate, and dynamic storage for updating the initialized weights.

63. The system of claim 51, wherein the analyzer initializes weights using random values approximately in the range of about – 0.5 to 0.5.

64. The system of claim 51, wherein the analyzer computes the normalized system error by using desired values and the predicted values, and the comparator compares the computed normalized system error with the normalized system error threshold value to reduce the error in the weight space using a gradient descent technique based on the outcome of the comparison.

65. The system of claim 64, wherein the gradient descent technique uses the equation:

$$W_{ij}\,(n+1) \;=\; W_{ji}\,(n) + \eta\,(\delta_j \;\, o_i)$$

Wherein W_{ij} are the weights in a space of i rows and j columns, o_j is the actual output, δ_j is the desired output, and η is the learning rate.

66. The system of claim 65, wherein the analyzer enhances the learning rate of the neural network during updating of weights using a Hessian Matrix:

$$H[i,k] = \sum_k \sum_j G[k] * w[i,j]$$

wherein H[i, k] are diagonal elements of second order derivatives, wherein i, j, and k are an architecture dependent number of nodes and hidden layers, and wherein

$$G[k] = \sum_j w[j,k] * d[i,k]$$

wherein $G\,[k]$ is a gradient of the previous iteration error with respect to the weight space and i, j, and k are an architecture dependent number of nodes and hidden layers.

67. The system of claim 66, wherein the analyzer further enhances the learning rate of the neural network using a function approximation neighborhood technique to update weights using the equation:

$$W\,(t+1) = f(n,\ m,\ \delta r(t)\,)$$

wherein n is a number of nodes in the input layer, m is a number of nodes in the next layer, and $\delta r(t)$ is a parameter based on a function of time.

68. The system of claim 67, further comprising:

repeating the varying learning rate, performing, and comparing to reduce the normalized system error.

69. The system of claim 51, wherein the analyzer validates the neural network to verify reliability of the learnt neural network by performing a supervised learning using learnt neural network to predict values, wherein the analyzer computes the accuracy of the predicted values by comparing the predicted values with known values, wherein the comparator compares the computed accuracy with an accuracy value, and the analyzer repeats the supervised learning based on the outcome of the comparison by the comparator.

70. The system of claim 69, further comprising:

repeating the varying learning rate of the neural network using techniques selected from the group consisting of Hessian matrix and function approximation neighborhood.

71. The system of claim 70, further comprising:

repeating the performing and validating of the neural network to enhance reliability of the learnt neural network.

72. The system of claim 71, further comprising:

an output module coupled to the analyzer to use the validated neural network to predict values.